UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P O Box 1450
Alexandria, Virginia 22313-1450
www uspto gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/536,739	05/27/2005	Jens Spille	PD020112	4699

24498 7590 01/27/2010
Robert D. Shedd, Patent Operations
THOMSON Licensing LLC
P.O. Box 5312
Princeton, NJ 08543-5312

EXAMINER	
LERNER, MARTIN	

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
01/27/2010	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/536,739	SPILLE ET AL.	
	Examiner	Art Unit	
	MARTIN LERNER	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *20 November 2009*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *10 to 14* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *10 to 14* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☒ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☒ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *9/30/2009*.

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 102

1. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

2. Claims 10, 12, and 13 are rejected under 35 U.S.C. 102(e) as being anticipated by *Lin et al. ('018).*

Regarding independent claim 10, *Lin et al. ('018)* discloses a method of encoding and decoding a presentation of audio data, comprising:

"transforming the 2D location information to a 3D coordinate system, wherein said y-location is mapped to audio depth information perpendicular to the 2D video plane and said x-location is mapped to itself" – video image 50 is shown containing two video objects 52, 54 that were previously extracted and matched with associated sound sources (e.g., sound source 1 and sound source 2); video object 52 is a person located in the lower right portion of the video image, and having a face located at column 6, row 3 of the two dimensional grid; video object 54 is a person located in the upper left hand portion of video image 50 and having a face located in column 1, row 1 of the two

dimensional grid (column 4, lines 30 to 55: Figure 2); in order to determine position data regarding a third dimension (i.e., depth), it is determined that video object 52 is closer to the viewer than video object 54; a size analysis system 40 could be used to determine the relative depth position of different objects in a three dimensional space based on the relative sizes of the video objects (column 4, line 56 to column 5, line 8: Figure 2); implicitly, then, the x and y coordinates remain the same for the audio data ("said x-location is mapped to itself"), but the size of the object, its relative y-location, determines the relative depth position of the audio object;

"adding a third coordinate value to the transformed location information in the 3D coordinate system; and spatializing the sound according to the resulting 3D location information" – the source associated with video object 52 can be assigned to a channel, or mix of channels, that would provide a sound image that is nearby the viewer, while the sound source associated with video object 54 could be assigned to a mix of audio channels that provide a distant sound image (column 4, line 56 to column 5, line 8: Figure 2); a system could be implemented that reconstructs a 3-D space based on the two dimensional video image 50; each sound source can then be assigned to an appropriate audio channel in order to create a realistic 3-D sound field ("spatializing the sound") (column 5, lines 9 to 29: Figure 2).

Regarding claim 12, *Lin et al. ('018)* discloses that a two dimensional video image 50 is located on a two dimensional grid comprising eight vertical columns and six horizontal rows (column 4, lines 40 to 55; column 6, lines 9 to 17: Figure 2); thus, the rows and columns correspond to "said x and y coordinates" of "the screen plane".

Regarding claim 13, *Lin et al. ('018)* discloses a size analysis system 40 could be used to determine the relative depth position of different objects in a three dimensional space based on the relative sizes of the video objects; the source associated with video object 52 can be assigned to a channel, or mix of channels, that would provide a sound image that is nearby the viewer, while the sound source associated with video object 54 could be assigned to a mix of audio channels that provide a distant sound image (column 4, line 56 to column 5, line 8: Figure 2); implicitly, a size of a video object changes as it moves closer to or farther away from the screen plane; motion analysis system 34 detects a video object in motion, *e.g.*, a moving car (column 4, lines 14 to 22); thus, movement of the object in a direction perpendicular to the screen plane produces at least a change in a vertical size of the object, and an apparent change in the sound image of the object as being nearer or more distant from the viewer in a direction perpendicular to the screen plane follows from the change in size.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. Claims 11 and 14 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Lin et al. ('018)* in view of *Scheirer et al. ("AudioBIFS: Describing Audio Scenes with the MPEG-4 Multimedia Standard").*

Lin et al. ('018) notes application to MPEG-4 (column 3, lines 31 to 33), and coding separate sound sources as separate audio objects (column 3, lines 22 to 25), but does not expressly disclose sound sources are described by a parametric scene description having a hierarchical graph structure with nodes, wherein a first node comprises x-location and y-location information, and a second node describes a third coordinate value and data describing the transformation. Moreover, *Lin et al. ('018)* omits mapping by a 2x3 vector or corresponding rotation.

However, it is known to represent sound sources as first nodes and presentation characteristics of sound sources as second nodes in MPEG-4 as taught by *Scheirer et al.* Specifically, *Scheirer et al.* teaches that AudioBIFS in MPEG-4 represent sound scenes, where an AudioClip node provides audio data that can be referenced by Sound nodes. A hierarchical audio subgraph represents each "child" node as presenting output resulting from one or more "parent" nodes. (III. A. The MPEG-4 Audio System: Page 242: Left Column: Figure 3) An AudioClip can be thought of as a property of the Sound node. The Sound node specifies the location (spatial position) of a sound object in a VRML scene, and a spatialize field specifies whether or not the audio object will be spatialized when presented. (II. C. Sound Scenes in VRML: Pages 238 to 240: Figures 1 and 3) Moreover, 3-D spatialization can be performed according to sound location in the corresponding azimuth and elevation angles. (III. B AudioBIFS Nodes: Page 244, Right Column) Implicitly, movement of an object in spherical coordinates corresponds to a rotation of the azimuth and elevation angles. An objective is to enable concise transmission of audiovisual scenes, and to provide a unified framework for sound

scenes that use streaming audio and three-dimensional (3-D) spatialization. (I. Introduction: Page 237) It would have been obvious to one having ordinary skill in the art to represent sound sources in a hierarchical graph structure with nodes corresponding to audio objects and presentation of audio objects includes spatialization as taught by *Scheirer et al.* in an audio encoding and decoding system of *Lin et al. ('018)* for a purpose of enabling concise transmission and a unified framework of sound scene spatialization in MPEG-4.

Response to Arguments

5. Applicants' arguments filed 20 November 2009 are sufficient to overcome the new matter rejection of claims 10 to 14 under 35 U.S.C. §112, 1st ¶.

6. Applicants' arguments filed 20 November 2009 directed to the rejections of claims 10, 12, and 13 under 35 U.S.C. §102(e) as being anticipated by *Lin et al. ('018)*, and of claims 11 and 14 as obvious under 35 U.S.C. §103(a) over *Lin et al. ('018)* in view of *Scheirer et al. ("AudioBIFS")*, have been fully considered but they are not persuasive.

Applicants' comments directed to how the transformation is performed employing the notation of the transformation mapping $\{x_i, y_i\} \rightarrow \{x_i, 0, y_i\}$ is appreciated, and is sufficient to overcome the new matter rejection under 35 U.S.C. §112, 1st ¶, for claims 10 to 13. Additionally, although Applicants did not argue the new matter rejection of claim 14 for the limitation of a "corresponding rotation", the Specification, Page 8, Lines 12 to 14, does disclose a "field data type 'SFRotation'", and so may be taken to

reasonably suggest a field dimension mapping that could equivalently be achieved by 'SFRotation'.

However, Applicants' arguments directed to the rejection of claims 10, 12, and 13 under 35 U.S.C. §102(e) as being anticipated by *Lin et al. ('018)* are not persuasive. Firstly, Applicants argue that *Lin et al. ('018)* describes a sound imaging system that creates position enhancement data relating to visible objects based on image analysis, and which results in depth information. Applicants say that *Lin et al. ('018)* processes mono audio data by adding a depth component according to the depth information, and outputs multi-channel audio. Applicants attempt to contrast *Lin et al. ('018)* with their method by stating that their spatialization is restricted to 2D audio input, whereas *Lin et al. ('018)* adds depth information by image analysis. Using their own notation, Applicants say that *Lin et al. ('018)* enhances a multi-channel audio signal $\{x=x_i, y=y_i\}$ of a 2D space to a 3D audio signal of the form $\{x=x_i, y=y_i, z=z_i\}$. Applicants maintain that because the added depth information z_i is obtained from image analysis, *i.e.*, from the size of an object, which can be designated as $z_i = f(x_i,y_i)$, *Lin et al. ('018)* must disclose a 3D audio signal that is enhanced as $\{x=x_i, y=y_i, f(x_i,y_i)\}$. Applicants conclude by saying that this transformation is different from their 3D coordinate system transformation.

The problem with Applicants' argument is that the independent claim is not limited to any specific transformation system, and *Lin et al. ('018)* appears to fully meet the terms of the claim as it is. Applicants might try to consider amending independent claim 10 to adapt it with further limitations of their parametric description. Although the Specification provides some evidence of a transformation mapping $\{x_i, y_i\} \rightarrow \{x_i, 0, y_i\}$,

there does not appear to be any express disclosure of a transformation mapping $\{x_i, y_i\} \rightarrow \{x=x_i, y=c_i, z=y_i\}$, which appears only in Applicants' arguments. The new second coordinate, c_i, does not appear to be clearly disclosed, except when it is zero.

Moreover, even supposing that *Lin et al. ('018)* discloses extracting the depth dimension of the audio object by image analysis, and that Applicants' method does not employ image analysis, it is not apparent that the claims would distinguish over *Lin et al. ('018)*. In fact, Applicants' independent claim 10 refers to "the 2D video plane", says that the method is for spatialization of sound "relating to a video", and that the x-location and y-location are "corresponding to the x and y coordinates of the video", so that it is not clear how Applicants would spatialize the sound without depending upon a description of the object in a video coordinate space. *Lin et al. ('018)* discloses that object position system 30 employs 3D location system 38 and size analysis 40 to determine a 3-D sound source location. (Column 4, Line 30 to Column 5, Line 8: Figure 1) Thus, even if the size of the object is analyzed according to the function, $f(x_i,y_i)$, to determine the depth of the audio sound source, it does not appear that independent claim 10 distinguishes over *Lin et al. ('018)*. The size of the object relates to its height, $f(y_i)$, – which Applicants are calling the y-location. It may be true that the size of an object relates to its width, $f(x_i)$, too, but *Lin et al. ('018)* determines the depth of the audio sound source by at least its height. Indeed, the relative size of an object as a function of how far away it is can relate to either its height or its width, so that the x and y coordinates are interchangeable. It follows that as long as *Lin et al. ('018)* determines the depth of the sound source by at least its height, then *Lin et al. ('018)* will be adding a

third coordinate value to obtain a 3D coordinate system, where the sound is spatialized according to the resulting 3D location information by a mapping of the y-location, as claimed.

Furthermore, it is not clear that *Lin et al. ('018)* depends upon conventional image analysis, or that Applicants' method functions in a manner that is patentably different. Of course, *Lin et al. ('018)* does not employ Applicants' notation to describe the coordinate system, but one skilled in the art could readily see from the two dimensional grid of Figure 2, and the disclosure of vertical columns and horizontal rows, that *Lin et al. ('018)* is contemplating an x, y, z coordinate system. The fact that *Lin et al. ('018)* may be starting with mono audio data for each of the sound sources, and then providing a three dimensional sound field for a plurality of sound sources, does not appear to produce a patentable distinction. Nor does the fact that *Lin et al. ('018)* may utilize image analysis. Actually, *Lin et al. ('018)*, at Column 5, Lines 13 to 16, discloses that "it should be recognized that any system for locating video objects in a space, two dimensional or three dimensional, is within the scope of the invention." Thus, *Lin et al. ('018)* may not be limited to any image analysis by matching system 30. Admittedly, *Lin et al. ('018)* discloses a main embodiment that identifies individual sound source data objects from video objects through matching system 30 so as to identify the objects. After that, once the images are identified, the relative sizes of similar identified objects are used to determine the audio depth of the object (*e.g.*, people, automobiles, or dogs) in a 3-D space.

But neither does Applicants' Specification, or their claims, provide substantial detail on how the y-location is mapped to the audio depth information. Certainly, there might be some computation involved in calculating an object's audio depth from its relative y-location in comparison to similar objects by *Lin et al. ('018)*, where the details of the algorithm for calculating the relative size are somewhat sketchy. But, correspondingly, Applicants' disclosed method is even less descriptive in demonstrating precisely how the audio depth information is produced from the y-location. One skilled in the art can only conclude that the y-location is simply related to the size of the object, as in *Lin et al. ('018).*

Secondly, Applicants set forth arguments purporting to show the problems with *Lin et al. ('018)*, and how those problems are solved by Applicants' claimed method. Applicants say that *Lin et al. ('018)* does not describe how to obtain the actual depth value from the height value, and maintains that some scaling must be applied to obtain this depth information. Applicants argue that there are problems for *Lin et al. ('018)* when visual objects appear at the horizon, or at the edge of the screen, and that the system would only be applicable to objects that are not hidden behind other objects, or are not outside the screen. Moreover, Applicants state that the human ear is more responsive to horizontal audio information than height information, so that errors in audio depth are more disturbing than errors in audio height. Finally, Applicants say that the invention maps two existing audio dimensions to the horizontal audio space, which is reproduced more exactly than audio height. These arguments are not persuasive.

Basically, all of these purported advantages relate to features that are not expressly claimed, and so that any advantages would not generally be applicable to an anticipation rejection under 35 U.S.C. §102(e). Even if there is a problem for objects on the horizon, for objects at the edge of a field, or for objects that are hidden in *Lin et al. ('018)*, there is nothing in Applicants' independent claim 10 that expressly provides any patentably distinguishing feature. Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993). Even more significantly, it is not understood that Applicants' Specification addresses these issues any better than *Lin et al. ('018)*. One might speculate that the locations of the video objects in a three dimensional virtual space may be initially given in Applicants' method, rather than having to deduce the depth dimension from given standard two dimensional video images by image analysis as might be the situation for *Lin et al. ('018)*. Similarly, it might be true that the human ear has better left-right audio perception than up-down audio perception. Still, it is not seen how these advantages relate to any claimed feature. Applicants' claims only perform the transformation from the y-location. However, it is not understood how the y-location relates to anything rather than indirectly being an indication of the size of the object for purposes of locating the audio depth information in Applicants' method. While something that could be called 'object scaling' may be used to obtain the relative size and, then, the depth of the object for *Lin et al. ('018)*, Applicants' method of obtaining the audio depth from only the y-location of the object, if anything, has a less straightforward description.

Applicants could amend their claims to more narrowly recite the nature of the sound field description, as disclosed by the Specification, Page 4, Lines 5 to 30. The exemplary embodiment there defines a 2D sound node by four variables: intensity, location, source, and spatialize, and a 3D sound node by ten variables: direction intensity, location, maxBack, maxFront, minBack, minFront, priority, source, and spatialize. Thus, although any amendment might require a new search, it is conceivable that Applicants could narrow their claims to somehow clearly distinguish over *Lin et al. ('018)*.

Therefore, the rejections of claims 10, 12, and 13 under 35 U.S.C. §102(e) as being anticipated by *Lin et al. ('018)*, and of claims 11 and 14 under 35 U.S.C. §103(a) as being unpatentable over *Lin et al. ('018)* in view of *Scheirer et al. ("AudioBIFS: Describing Audio Scenes with the MPEG-4 Multimedia Standard")*, are proper.

Conclusion

7. **THIS ACTION IS MADE FINAL.** Applicants are reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of

the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to MARTIN LERNER whose telephone number is (571)272-7608. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Martin Lerner/
Primary Examiner
Art Unit 2626

January 22, 2010